

XF-S3011 中文语音合成芯片

数据手册

V 1.4 – 2006 年 11 月 14 日

安徽中科大讯飞信息科技有限公司

安徽省合肥市国家级高新技术产业开发区信息产业基地讯飞语音大厦

86-0551- 5331878/5397818 86-0551- 5331837 www.iflytek.com

版本历史

版 本	日 期	修改记录	作 者
Rev1.0	2004-12-22	完成本手册第一版	科大讯飞
Rev1.1	2005-5-11	强调音频电路设计的注意事项	科大讯飞
Rev1.2	2005-5-18	修改外围应用电路	科大讯飞
Rev1.3	2006-8-10	修改外围应用电路 RXD/TXD 部分	科大讯飞
Rev1.4	2006-11-14	修改联系方式	科大讯飞

声 明

科大讯飞保证本手册提供信息的准确性和可靠性，但并不对文本中可能出现的文字错误或疏漏负责。科大讯飞保留更改本手册的权利，如有修改，恕不相告。请在订购时联系我们以获得产品最新信息。对任何用户使用我们产品时侵犯第三方版权或其他权利的行为科大讯飞概不负责。另外，在讯飞未明确表示产品有该项用途时，对于产品使用在极端条件下导致一些可预见失灵或损毁而造成的损失概不负责。

目 录

1. 概述

语音合成技术，简称 TTS（Text To Speech）技术，该技术解决如何将文字信息转化为声音信息的问题，从而变看为听，使得人们获取信息的方式更加丰富和自然。语音合成技术是实现人机语音通信的关键技术之一，涉及声学、语言学、数字信号处理、计算机科学等多种学科知识，是中文信息处理领域的一项前沿技术。

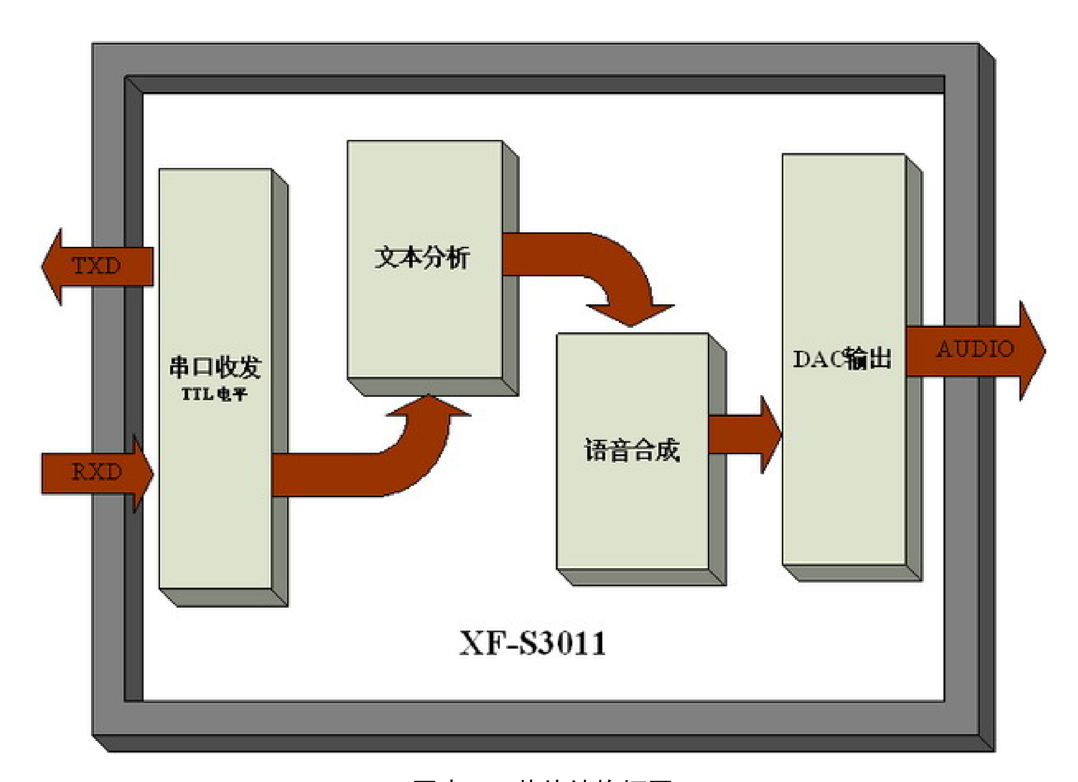
XF-S3011 芯片是安徽中科大讯飞信息科技有限公司研发的一款中文语音合成芯片，将完整的语音合成系统完全集成到单一的处理器内部，通过串口接收并合成任意文本。该芯片采用QFP*封装，可方便的集成到需要中文语音合成功能的嵌入式设备之中。

2. 特点

- 可合成任意的中文文本，支持英文字母的合成
- 支持 GB2312、UNICODE 两种内码格式的文本
- 具有智能的文本分析处理算法，可正确的识别和处理数值、号码、时间日期及一些常用的度量衡符号，具备一定的多音字处理和中文姓氏处理能力
- 清晰、自然、准确的中文语音合成效果
- 对电话号码、数值、时间日期等特定文本提供较好的定制合成效果
- 芯片内固化有多首和弦音乐、提示音效和针对某些行业领域的常见语音提示音
- 支持多种控制命令，包括：合成、停止、暂停合成、继续合成等
- 支持多种文本控制标记，提升文本处理的正确率
- 支持休眠功能，在休眠状态下可降低功耗
- 支持 UART 串行数据通讯接口，波特率 9600bps
- 两路 10 bit DAC 模拟音频输出
- 宽工作电压：2.6~5.5V
- 提供 QFP（44pin）封装形式和裸片封装形式
- 芯片各项指标均满足车载环境下的应用

QFP：即 Quad Flat Package 详情可查询有关标准规定。

3. 系统框图



图表 1. 芯片结构框图

4. 主要应用领域

- ✓ 车载 GPS 导航
- ✓ 车载电话、信息电话
- ✓ 公交车语音报站器
- ✓ 考勤机、打卡机
- ✓ 税控机、POS 机
- ✓ 智能仪器
- ✓ 智能玩具
- ✓ 排队机
- ✓ 自动售货机

5. 功能描述

XF-S3011 中文语音合成芯片是一款针对嵌入式应用领域而设计的单芯片处理器，具有合成任意中文文本的能力。芯片的文本分析算法具备一定的智能性，可识别常见的数值、号码、时间、度量单位等格式的文本，可对中文姓氏中的多音字进行处理。除了基本的语音合成功能以外，芯片内还定制了一些和弦音乐、提示音效和少数特定行业的常用语音提示音，使得芯片的效果有更佳表现；同时芯片提供特殊的控制标记，使用这些标记可以让合成的效果达到更好的效果。由于采用小尺寸的单芯片封装架构，用户可以方便的将其集成到嵌入式设备中，为设备添加人性化的语音功能。该芯片的主要功能描述如下：

- 文本合成功能：芯片支持任意中文文本的合成，可以采用 GB2312 或者 Unicode 两种编码方式，其中 GB2312 是中国大陆地区常用的中文文本编码格式；Unicode 编码是国际通用的编码格式。芯片在默认情况下按 GB2312 编码格式处理中文文本。芯片支持英文字母的合成，遇到英文单词时按字母方式发音。

- 文本分析算法：芯片具有智能的文本分析算法，对常见的数值、电话号码、时间日期、度量衡符号等格式的文本，系统能够根据内置的文本匹配规则进行正确的识别和处理。如“2004/12/21”可以识别并读作 “二零零四年十二月二十一日”。

- 定制合成引擎：芯片对号码、数值、时间日期类型的文本，定制了特殊的处理引擎。如果采用定制引擎合成上述文本，可以获得更加自然流畅的文本合成效果，尤其适用于经常要播报数值、号码和时间的场合。

- 提示音：芯片内有和弦提示音、音效提示音，可用于铃声和信息提醒。此外，还针对车载电话、车载导航、考勤机、智能仪器、税控机等行业，录制了一些常用的语音提示音，配合定制合成引擎的使用，可以使这些行业领域的信息播报获得更为自然的效果。

- 文本标记功能：芯片允许在文本中插入特殊控制标记，利用这些控制标记可以调节芯片的参数，提升芯片文本分析的正确率，如：可调节芯片合成的音量大小、产生静音间隔、标识中文姓氏文本、标识数值、号码文本等。

- 简易的开发接口：芯片采用串口通信方式，且命令协议简练有效；可通过统一的“合成命令”接口播放音乐、音效和中文文本，甚至进行系统参数的设置。开发者只需要掌握少量命令便可轻松上手，使用所有主要功能。

- 硬件特性：芯片采用单芯片封装形式，尺寸小，易于集成；具有休眠功能，可节约能耗；可通过引脚电平查询芯片当前的工作状态。

6. 通讯方式

XF-S3011 通过串口(UART)接收上位机发送的命令和数据，命令和数据以帧的方式进行封装和发送，允许发送数据的最大长度为 200 个字节。芯片的电平逻辑为 TTL 电平，

6.1 通讯波特率

9600bps

6.2 通讯传输字节格式



图表 2. 通讯传输字节格式

串口数据传输字节格式：1 bit 起始位，8 bit 数据位，1 bit 停止位，没有校验位。

7. 控制命令

7.1 控制命令描述

XF-S3011 芯片可以接受多种控制命令。控制命令包括以下几种：

控制命令	说明
合成命令	合成本次发送的文本
停止合成命令	停止当前的合成动作
暂停合成命令	暂停正在进行的合成
恢复合成命令	继续合成被暂停的文本
带文本编码设置的合成命令	以指定的编码格式(GB2312 或 Unicode),合成本次发送的文本
休眠命令	芯片进入休眠节能状态，如要唤醒需要重新复位芯片

图表 3. 控制命令列表

注：当芯片正在合成文本的时候，如果又接收到一帧有效的合成命令，芯片会立即停止当前正在合成的文本，转而合成最新接收到的文本。

7.2 芯片反馈

XF-S3011 芯片在接收到命令或文本合成完毕之后，通过串口向上位机发送反馈信息。上位机可以通过接收芯片的反馈，了解到芯片是否正确接收了命令，以及芯片是否完成了文本合成。

7.3 芯片状态查询

芯片平时处于待机状态，当收到合成命令后，会进入合成状态，待合成结束，又会恢复到待机状态。如果用户想得知芯片当前的工作状态，除了可以通过串口接收芯片发送的反馈以外，还可以通过芯片的$\overline{RDY}$或RDY引脚的电平状态来进行查询。

当$\overline{RDY}$引脚处于低电平时，表明芯片处于待机就绪状态；

当$\overline{RDY}$引脚处于高电平时，表明芯片正处于文本合成状态。

而RDY引脚信号与$\overline{RDY}$引脚信号相反。用户可以根据情况任选其中一根引脚来实现对芯片的状态查询。

7.4 芯片的复位

芯片的复位引脚是低电平有效，其复位时序图如下所示：

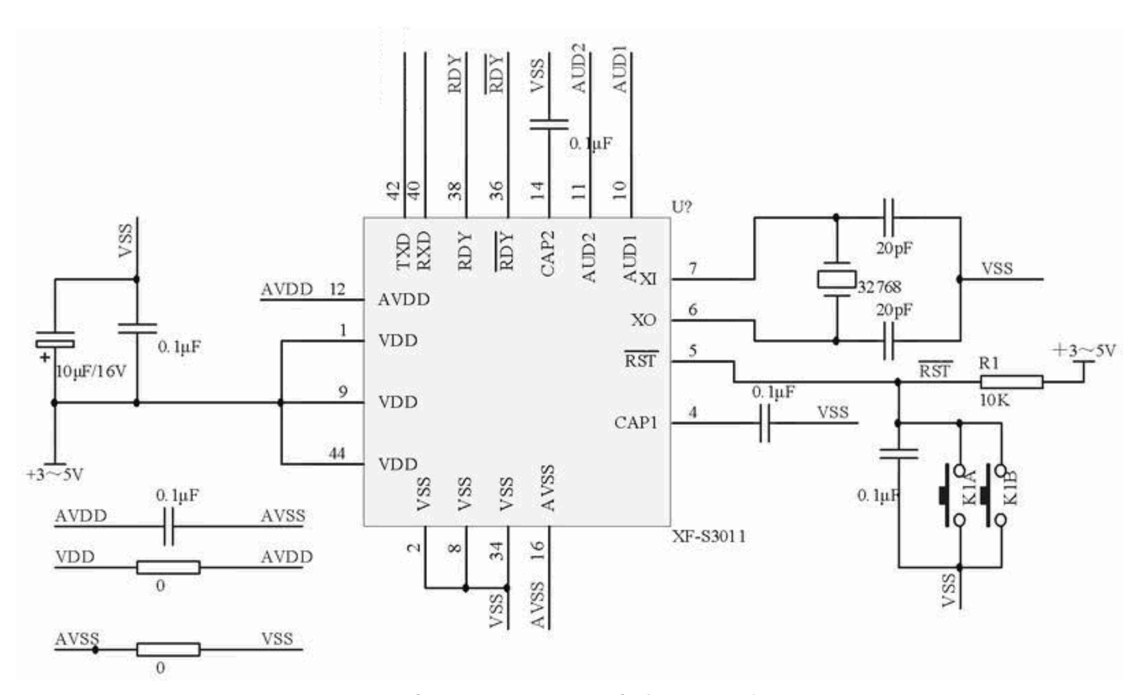


图表 4. 芯片复位

注：T_{RST}的典型时间是 10ms；当T_{RST}上升到高电平以后，需要等到 50ms以后才能开始给芯片发送命令/数据，以保证芯片可靠复位。

8. 外围应用电路

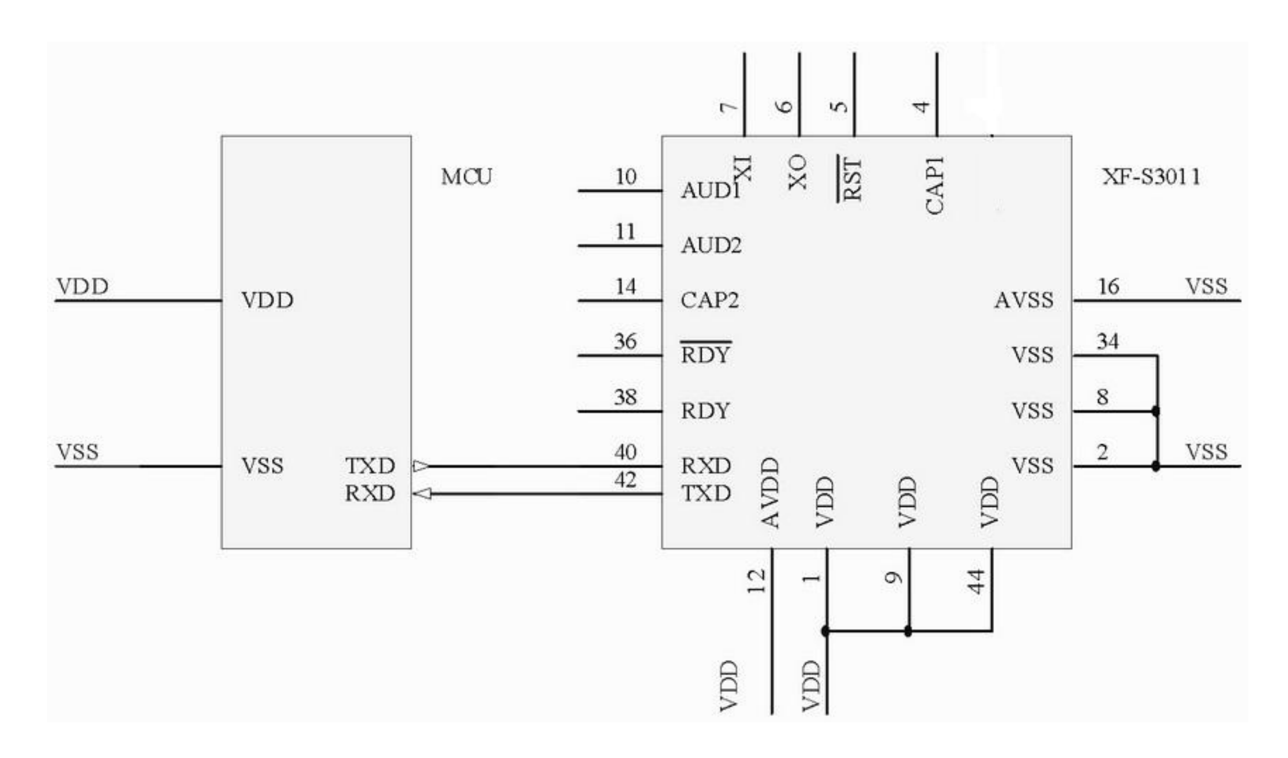
8.1 芯片外围电路



图表 5. XF-S3011 芯片外围电路

8.2 芯片以单片机为上位机的连接

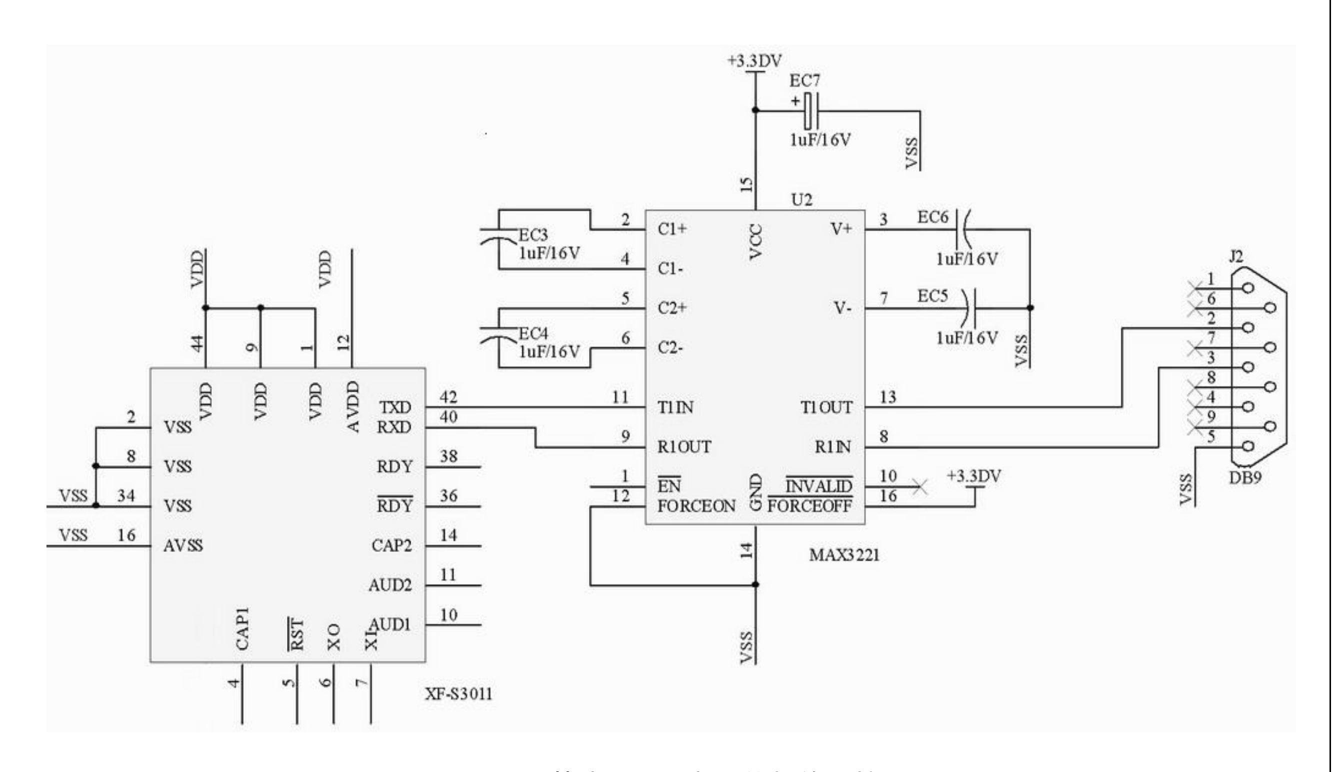
芯片与单片机（单片机用户自行选定型号）连接，将单片机的 TXD 与芯片的 RXD 相连，单片机的 RXD 与芯片的 TXD 相连，用于双方通讯。



图表 6. XF-S3011 芯片以单片机为上位机的连接

8.3 芯片以 PC 为上位机的连接

芯片与 PC 连接，由于 PC 串口为 RS232 电平而 XF-S3011 芯片为 TTL 电平，双方进行串口通讯时需要有电平转换。参考电路中采用的是 MAX3221，将 RS232 电平降至 2.7V，实现双方通讯。



图表 7. XF-S3011 芯片以 PC 为上位机的连接

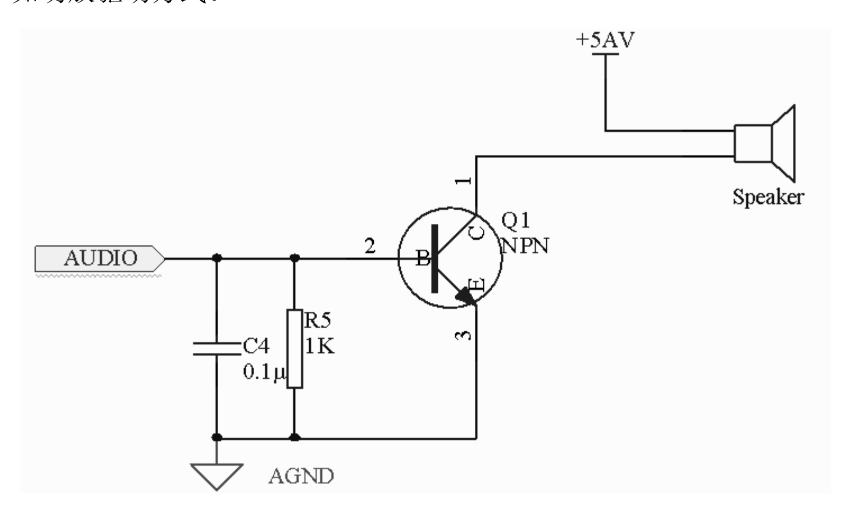
8.4 音频参考电路

XF-S3011 芯片有两路 10bitDAC 模拟音频输出，AUD1 和 AUD2 的参考电路图是一致的，用户可以根据自己的需要选择一条音频输出使用，也可以同时使用两条音频输出。音频信号可直接用三极管驱动方式放大，也可采用功放电路放大。

注：音频电路的质量对最终语音合成效果的影响较大，因此建议采用本节中的音频参考电路进行设计。如果用户使用自己的播音电路，那么应该注意检查是否实现了必要的滤波电路，是否保证正确接地等，否则会容易出现音量较小或者噪音较大的问题。

8.4.1 三极管驱动

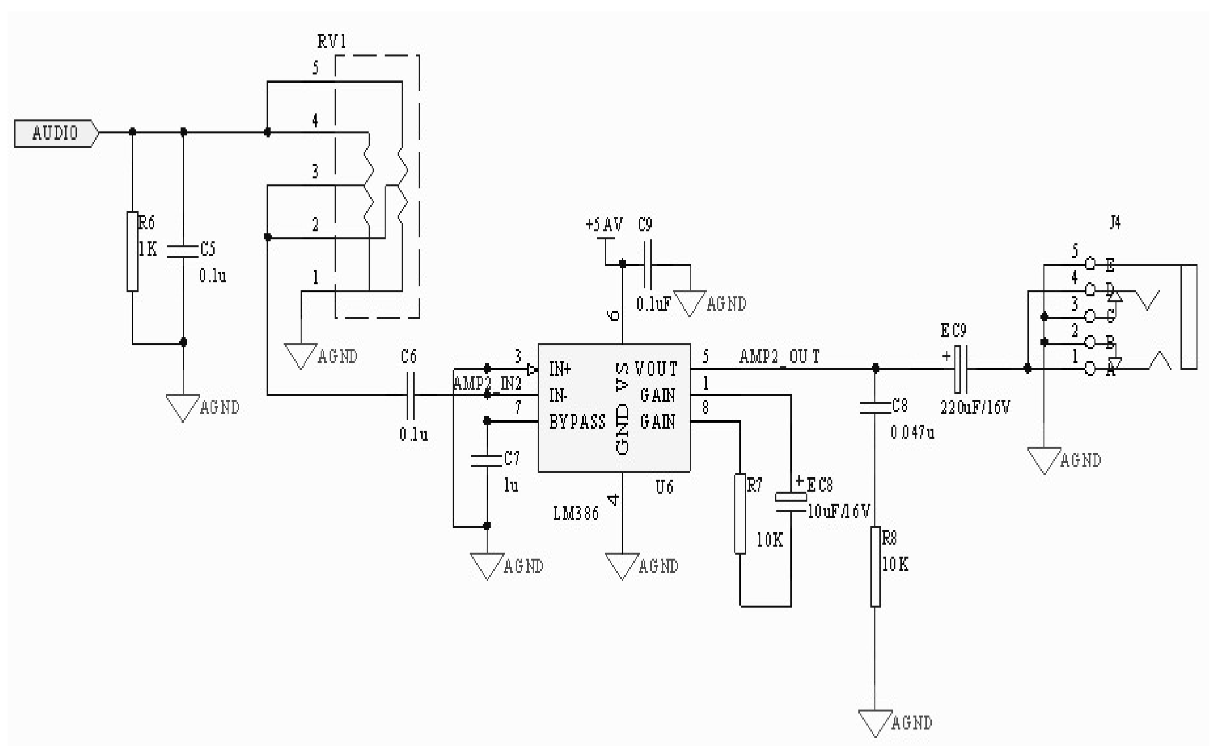
三极管方式驱动电路简单，成本较低，但噪声比较大，增益较小，因此输出音量小，音质不如功放驱动方式。



图表 8. XF-S3011 芯片音频电路（三极管驱动）

8.4.2 功放驱动

功放方式驱动电路较三极管方式复杂，成本较高，但可获得较大增益，因此输出音量大，音质更高。



图表 9. XF-S3011 芯片音频电路（功放输出）

9. 芯片封装形式

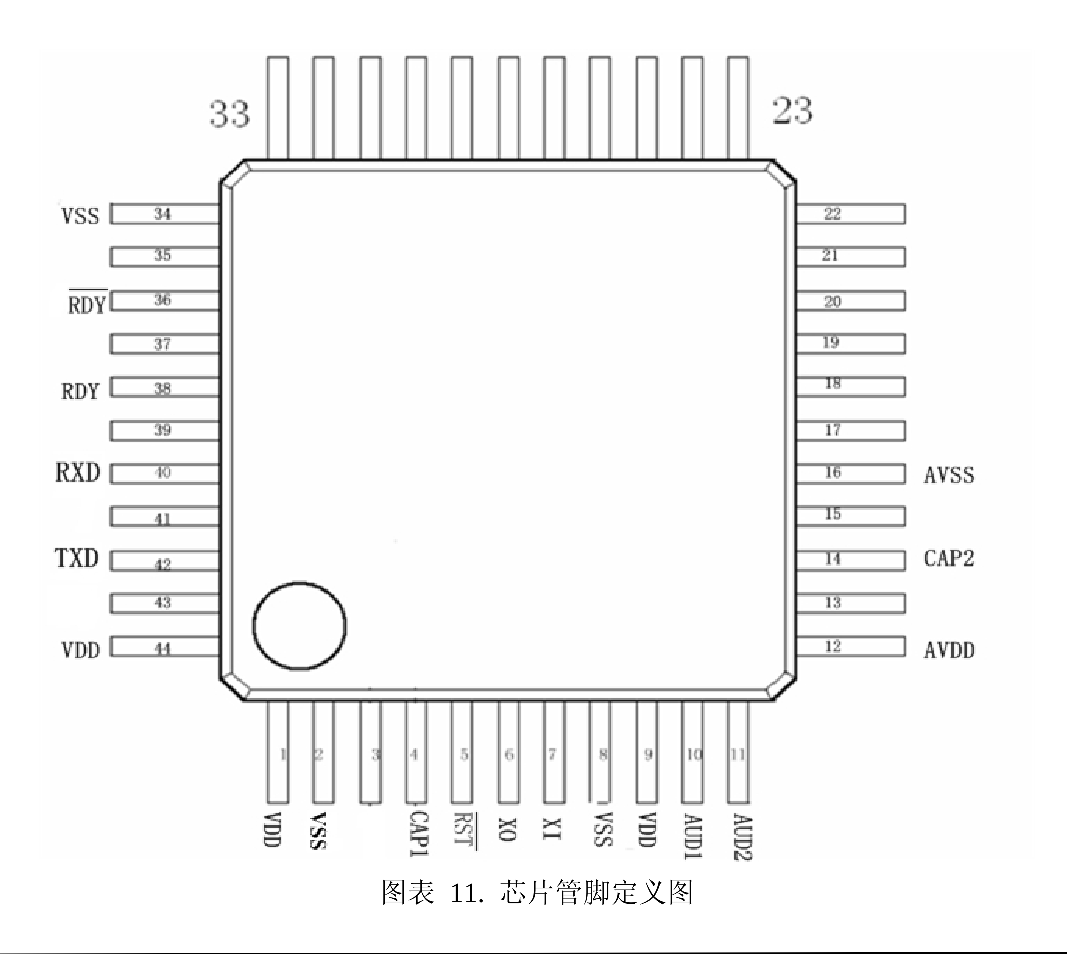
9.1 外观

XF-S3011 芯片采用 QFP44-2 的封装形式，下图是实物照片：



图表 10. 芯片 QFP 封装外观图

9.2 引脚定义



图表 11. 芯片管脚定义图

引脚序号	信号名称	说明
1	VDD	数字电源输入
2	VSS	数字电源地
4	CAP1	外接滤波电容
5	$\overline{RST}$	芯片复位引脚，低电平有效
6	XO	晶振输出
7	XI	晶振输入
8	VSS	数字电源地
9	VDD	数字电源输入
10	AUD1	模拟音频信号输出 1
11	AUD2	模拟音频信号输出 2
12	AVDD	模拟电源输入
14	CAP2	外接滤波电容
16	AVSS	模拟电源地
34	VSS	数字电源地
36	$\overline{RDY}$	芯片空闲指示，低电平有效
38	RDY	芯片空闲指示，高电平有效
40	RXD	芯片的串口接收脚
42	TXD	芯片的串口发送脚
44	VDD	数字电源输入
3,13,15,17~33,35,37,41,43	NC	Not Connect

图表 12. 芯片引脚定义列表

注：

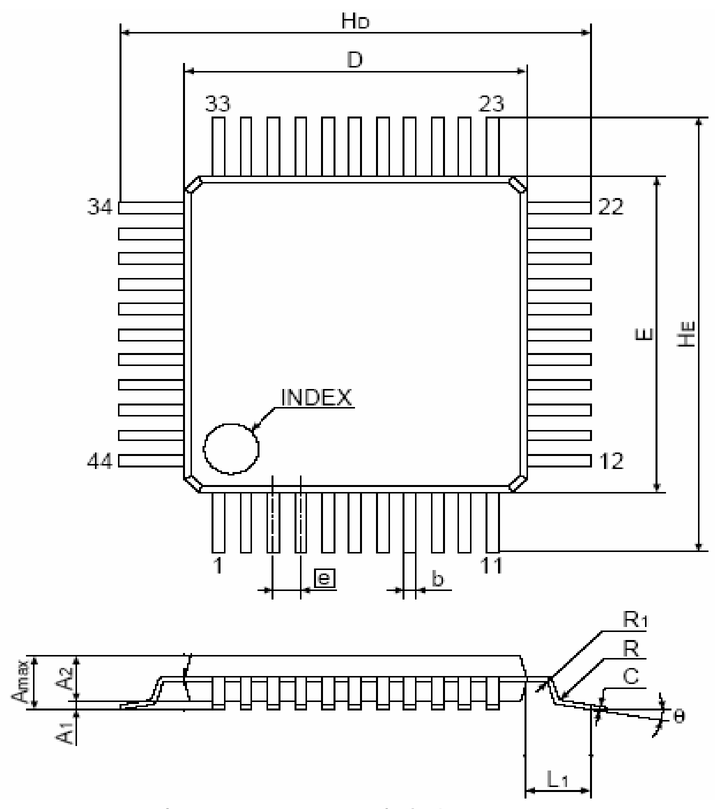
1．VDD:2.6V～5.5V；

2．芯片的工作状态指示信号有两根，一是高电平有效，一是低电平有效，您可以根据自己的实际情况选用其中的一根；

3．芯片的 40 脚为 RXD，它和上位控制机的数据发送端相连；42 脚为 TXD，它和上位控制机的数据接收端相连；

4．芯片有两路 10bit DAC 模拟音频输出：AUD1、AUD2，客户可以任选其中一根进行音频输出。

9.3 装配图



图表 13. XF-S3011 芯片装配图

Symbol	Dimension（mm）
E	10
D	10
A_1	0．14
A_2	2
Amax	2．14
e	0．8
C	0．25
θ	5°
b	0．3
R	0．30
R_1	0．15
L_1	1．85
H_E	13．5
H_D	13．5

10. 电气特性

10.1 极限参数

参　数	符　号	极　限　值
电源电压	VDD	5.5V
引脚输入电压范围	V_{IN}	－0.5V to VDD+0.5V
工作温度	T_A	－35℃ to +85℃
储存温度	T_{STO}	－40℃ to +100℃

图表 14. 电气极限参数

注意：如果超出表中所列的极限参数，将可能导致器件损坏。

10.2 直流特性

直流特性（VDD ＝ 3.3V TA ＝ 25℃）

参　数	符 号	参　考　值			单位
		最小	典型	最大	
工作电压	VDD	2.6	5.0	5.5	V
工作电流	I_{OP}	--	12	--	mA
输入高电平	V_{IH}	0.7VDD	--	--	V
输入低电平	V_{IL}	--	--	0.3VDD	V
输出电流	I_{OH}	--	-1.0	--	mA
吸收电流	I_{OL}	--	4.0	--	mA
DAC 输出电流	I_{DAC}	--	-1.5	--	mA

直流特性（VDD ＝ 5.0V TA ＝ 25℃）

参　数	符 号	参　考　值			单位
		最小	典型	最大	
工作电压	VDD	2.6	5.0	5.5	V
工作电流	I_{OP}	--	20	--	mA
输入高电平	V_{IH}	0.7VDD	--	--	V
输入低电平	V_{IL}	--	--	0.3VDD	V
输出电流	I_{OH}	--	-1.0	--	mA
吸收电流	I_{OL}	--	4.0	--	mA
DAC 输出电流	I_{DAC}	--	-3.0	--	mA

图表 15. 芯片直流特性